

## A. Overview Information

- I. Federal Agency Name: Defense Advanced Research Projects Agency, Information Processing Technology Office (DARPA/IPTO)
- II. Title: Learning by Reading (LBR)
- III. Announcement Type: Request for Information (RFI)
- IV. Solicitation Number: RFI 08-11
- V. CFDA Number: 12.910
- VI. Key Dates
  - a. Responses Due: any time from publication up to 4:00 PM (ET), February 1, 2008

## B. Full text of announcement

### I. Description

Today we have systems with extensive formal reasoning capabilities. In addition, information in the form of electronic text is becoming increasingly available. Unfortunately, today's formal reasoning systems cannot operate with a textual input. Thus, the purpose of this RFI is to explore general purpose bridges between the corpus of electronic text and formal reasoning systems.

The ability to perform a deep-level semantic analysis on the enormous available body of human knowledge that is available electronically would provide significant advantages in heightening situation awareness and recognizing and responding to asymmetric threats. Thus, capturing natural text as formal knowledge would, by itself, constitute a fundamental advancement that would afford extensive military benefits.

In accordance with FAR 35.007(j), the Information Processing Technology Office (IPTO), Defense Advanced Research Projects Agency (DARPA) requests information in the following area:

**Automated systems that accept natural language (English) text as input and provides some kind of formal, usable output. Possible output behaviors include, but are not limited to, knowledge base generation/augmentation, answering questions, and solving problems.**

The requested information is sought to identify the key technologies, concepts and components required to achieve the envisioned goals and promote productive interactions among the various relevant communities that must be combined to create the desired results. A Broad Agency Announcement and/or other solicitation may be posted at a later date, but the Government is not obligated to do so, and this RFI does not signify any firm intention by the Government to do so. In addition, DARPA/IPTO may hold a workshop on this topic at some point in the future. One is not currently

planned; however, in the event a workshop will be held, a notice will be posted at [www.fedbizopps.gov](http://www.fedbizopps.gov).

DARPA/IPTO is soliciting responses on the aforementioned topic. The response should provide answers to the following questions in the form of two PowerPoint slides and backup materials:

➤ **Slide 1 – Program Vision**

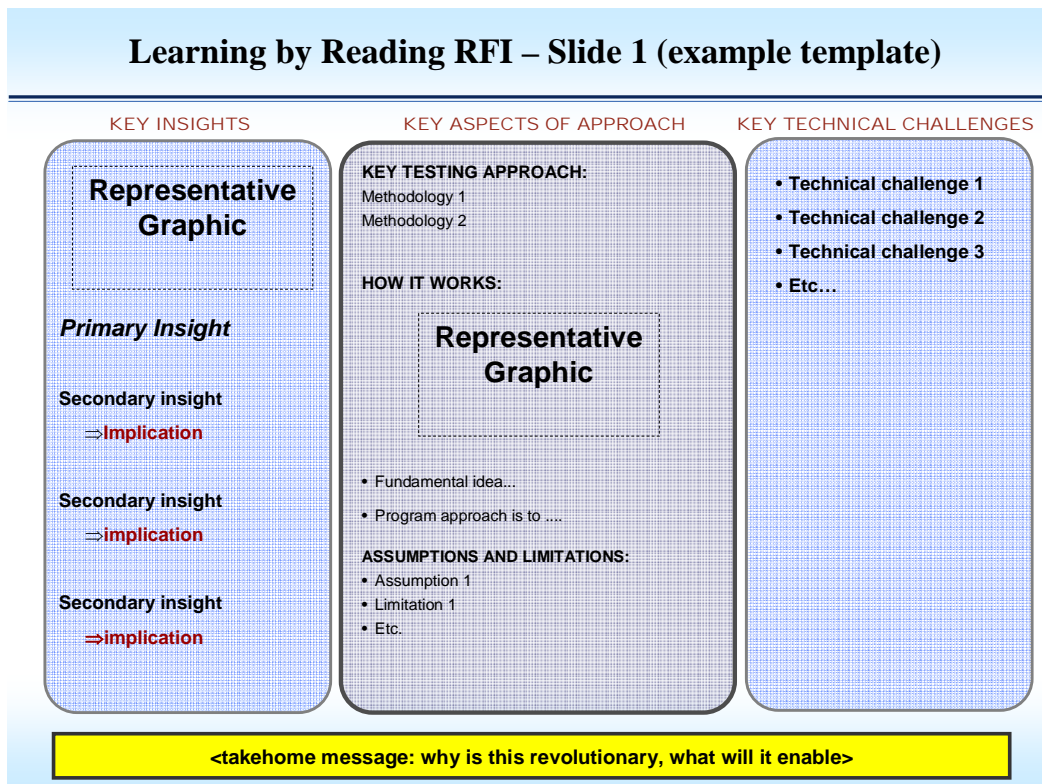
- What should be the goal and structure of a “Learning by Reading” program?
  - In one sentence, characterize the program goals: “This program will achieve the capability to...”
  - Provide a programmatic overview that describes succinctly:
    - new **key insights** that are germane to advancement
    - key aspects of approach
    - key technical challenge(s) to be overcome

➤ **Slide 2 – Programmatic approach**

- How should the program be constrained?
  - What are appropriate technical constraint dimensions?
  - What input, output, type of problem, and other constraints should define the program scope? (Obvious examples are: choice of reading level, sentence length, descriptive vs. transcriptive text, etc. What else should be considered?)
    - Syntactic/Semantic?
    - Problem types?
    - Analytic?
  - Is there a logical trajectory for relaxing these constraints over time as the system becomes more capable?
- What performance tasks and associated metrics should be used to evaluate success?
- What are the key programmatic issues?
- How should the program be partitioned into tractable stages? In other words, what is an appropriate sequence of program objectives?
- One possible goal for the program is to develop a system that receives English text as input and provides some useful knowledge representation as the output.
  - In such a case, what should the target representation or ontology be?
  - How would it be specified?
- Where applicable, what are the recommended data sources for training and testing the corpus?
- In order to delineate the programmatic goal, list problems that are reasonably in and out of scope of a “Learning by Reading” system.

**Slide 1 guidance:** This slide should present a vision of a possible Learning by Reading program that is both ambitious and attainable, drives important areas of research, whose testing is cost-effective (does not detract too much from research), and if successful, would be revolutionary. The Slide 1 Example below may serve as a guideline for organization and level of content. The format is not required and other formats are welcome.

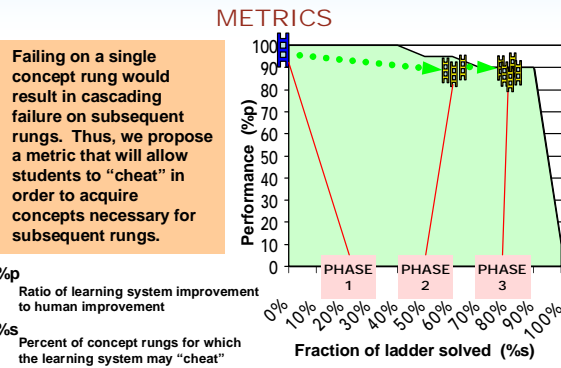
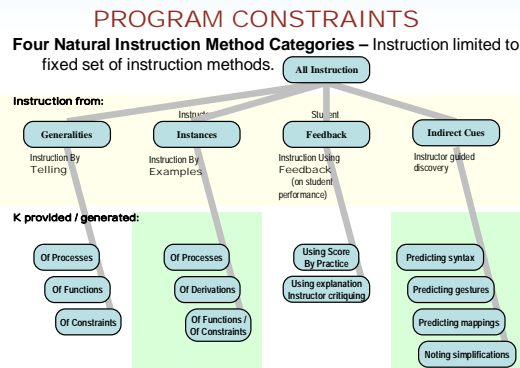
**Slide 2 guidance:** This slide should favor innovative ideas over completeness. It is much more desirable to see a few novel ideas presented than to have all of the questions answered. Furthermore, concepts should be presented as key phrases, which may be elaborated in the supportive materials to permit the reader to delve deeper. The expectation, however, is that slide content will be completely self-contained. Note: DARPA will focus primarily on the slides and may or may not read backup materials. Where elaboration occurs, there should be clear and obvious correspondence to associated slide content. The Slide 2 Example below depicts a partially filled-in slide (content has been drawn from an unrelated topical area) in order to illustrate the target conceptual level and graphical richness.



**Slide 1 Example**

## Learning by Reading RFI – Slide 2

(example template only; content drawn from another topic)



**<Answer to one of the other RFI questions>**

**<Answer to one of the other RFI questions>**

Slide 2 Example

Interested parties must submit their responses (two PowerPoint slides and backup materials) as described in Section IV - Application and Submission Information below. DARPA will acknowledge receipt of the submission, but will not provide feedback. However, in the event of an associated workshop, DARPA/IPTO might invite selected individuals to present talks and provide further discussion on relevant topics, based on their submissions. **Submissions must not contain any proprietary information.** All information submitted in response to this announcement will be considered public information and may be made available to workshop attendees. In the event of an associated solicitation, submissions may also be considered available to be placed on a public web site.

All information contained in the RFI is preliminary, as well as subject to modification, and is in no way binding on the Government. This RFI incorporates by reference FAR 52.215-3, "Request for Information or Solicitation for Planning Purposes (OCT 1997)," with the same force and effect as if it were given in full text [reference paragraph (c) of this provision, the "purpose" of this RFI is detailed in this announcement].

### II. Award Information

This notice, which constitutes the complete RFI package, is not a Request for Proposals (RFP), and is not to be construed as a commitment by the Government to issue a

solicitation or ultimately award a contract. Responses will not be considered as proposals nor will any award be made as a result of this synopsis. The Government is not interested in specific company capability information and will not entertain such submissions. Any costs incurred as a result of responding to this announcement shall be borne by the respondent and cannot be charged to the Government for reimbursement.

### III. Eligibility Information

#### 1. Eligible Applicants

All responsible sources capable of satisfying the Government's needs may submit a response that shall be considered by DARPA. Participation in this RFI is not required to propose to subsequent Broad Agency Announcements or research solicitations on this topic.

#### 2. Cost Sharing or Matching – N/A

### IV. Application and Submission Information

#### 1. Addresses to Request Application Package

This announcement contains all information required to submit a response. No additional forms, kits, or other materials are needed.

#### 2. Content and Form of Application Submission

DARPA/IPTO requires completion of an online RFI Cover Sheet for each response, by accessing the URL below:

<https://www.csc-ballston.com/rfi/rfiindex.asp?RFId=08-11>

After finalizing the Cover Sheet Submission, the offeror must submit the Confirmation Sheet that will automatically appear on the web page. Failure to comply with these submission procedures may result in the RFI response not being considered.

Submissions must be made via direct upload to DARPA. Instructions for doing so will be provided upon completion of the online cover sheet referenced above. Please note that submissions must be zipped together into a single file using PKZip or WinZip. Only one file can be uploaded, attempting to upload multiple files will result in the deletion of previously uploaded files. RFI responses are limited to the sections outlined below. Respondents are encouraged to be as succinct as possible while at the same time providing actionable insight. Each response should address the questions posed above and should comprise the following sections:

Section I. Cover Page (see Confirmation Sheet above) to include the following information:

- Respondent Name
- Organization
- Mailing address
- Phone number

- Fax number
- Email address

Section II. Two PowerPoint slides detailing the submitted position on research in the area of Learning by Reading, addressing the above questions and any others the submitter deems important. Respondents may also include up to twenty pages of supportive materials (in word or pdf format), understanding that the focus will be on the two slides and that the backup materials may not be read. All supportive material should be clearly indexed so that material corresponding to specific slide content can be easily retrieved.

Section III. Additional Optional Information: In addition to the formal submission, DARPA may or may not include other ad hoc forms of interaction (e.g., discussions, teleconferences, emails, workshops, etc.) in word or pdf format. If respondents are interested in being considered for participation in such interactions as this topic evolves, they may attach a curriculum vita of the principal or lead person representing the RFI submittal and a single paragraph of text (100 words or less) indicating the relevance of their background and the strength of their potential contributions. Including such a statement of interest does not guarantee future involvement.

### 3. Submission Dates and Times

Submissions will be accepted if they are received by 4:00 PM (ET), February 1, 2008. Responses may be sent at any time from publication up until the expiration date of February 1, 2008.

4. Intergovernmental Review – N/A
5. Funding Restrictions - N/A
6. Other Submission Requirements – N/A

### V. Application Review Information

1. Criteria – N/A
2. Review Process – N/A

### VI. Award Information Administration – N/A

1. Award Notices – N/A
2. Administrative and National Policy Requirements – N/A
3. Reporting – N/A

### VII. Agency Contacts

All administrative correspondence and questions concerning this announcement should be directed to one of the following administrative addresses:

Fax: 703-741-7804, Addressed to: DARPA/IPTO, Attn: RFI 08-11

Electronic Mail: RFI0811@darpa.mil

Electronic File Retrieval: <http://www.darpa.mil/ipto/Solicit/solicit.asp>